cDNA cloning of ovine pulmonary SP-A, SP-B, and SP-C: isolation of two different sequences for SP-B

SYLVIA M. PIETSCHMANN AND ULRICH PISON
Department of Anesthesiology and Intensive Care Medicine, Charité Campus Virchow-Klinikum, Medical Faculty Humboldt University Berlin, D-13353 Berlin, Germany

Pietschmann, Sylvia M., and Ulrich Pison. cDNA cloning of ovine pulmonary SP-A, SP-B, and SP-C: isolation of two different sequences for SP-B. Am J Physiol Lung Cell Mol Physiol 278: L765-L778, 2000.—Pulmonary surfactant promotes alveolar stability by lowering the surface tension at the air-liquid interface in the peripheral air spaces. The three surfactant proteins SP-A, SP-B, and SP-C contribute to dynamic surface properties involved during respiration. We have cloned and sequenced the complete cDNAs for ovine SP-A and SP-C and two distinct forms of ovine SP-B cDNAs. The nucleotide sequence of ovine SP-A cDNA consists of 1,901 bp and encodes a protein of 248 amino acids. Ovine SP-C cDNA contains 809 bp, predicting a protein of 190 amino acids. Ovine SP-B is encoded by two mRNA species, which differ by a 69-bp in-frame deletion in the region coding for the active airway protein. The larger SP-B cDNA comprises 1,660 bp, encoding a putative protein of 374 amino acids. With the sequences reported, a more complete analysis of surfactant regulation and the determination of their physiological function in vivo will be enabled.

THE EPITHELIUM OF THE LUNG synthesizes surfactant, a surface-active film that reduces surface tension at the air-liquid interface, thus preventing collapse of lung alveoli at the end of expiration. Pulmonary surfactant is a complex mixture of phospholipids, neutral lipids, and proteins. The surfactant-associated protein (SP) A, SP-B, and SP-C are known to play important roles in the biological functions of pulmonary surfactant. The most abundant SP-A is a hydrophilic glycoprotein that is involved in two major groups of function. First, it contributes to the surface tension-lowering capacity of pulmonary surfactant under certain in vivo situations. Second, it is involved in innate host defense and inflammatory processes of the lung (28, 38). The low-molecular-weight proteins SP-B and SP-C are strictly hydrophobic. Both SP-B and SP-C promote the formation of the surfactant film that floats on the alveolar lining layer in mammalian lungs (18, 23). The genes for SP-A, SP-B, and SP-C are transcribed and translated into much larger precursors from which the functional active SPs present in air spaces are generated by intracellular processing (29, 52).

A lack of functional active surfactant is recognized as the cause of the respiratory distress syndrome common in premature infants (10), and inactivation of surfactant contributes to the development of respiratory failure in adults (37). Several studies (7, 51) implied a primary role for SP-B in the metabolism of SP-A, SP-C, and surfactant phospholipids. SP-B also maintains the biophysical properties and physiological function of surfactant (8, 41, 42, 44). Deficiency of the active SP-B protein and/or its mRNA results in severe respiratory failure. Such deficiency is either inherited, as demonstrated in congenital alveolar proteinosis, the cause of respiratory failure in term neonates (31, 32), or the result of targeted disruption of the SP-B gene in transgenic mice (7). Furthermore, antibodies directed against SP-B induce severe respiratory failure in experimental animals (43, 48). Deficiency of SP-B also affects at least another SP because the processing and secretion of the SP-C precursor protein is aberrant in humans lacking the SP-B protein (51) or in SP-B-deficient mice (7).

Although sheep have been extensively used for studying pulmonary surfactant in health and disease (21, 49), the sequences for the ovine pulmonary SPs have not been provided. As a first step in studying gene regulation in a sheep model of lung injury, we wanted to characterize ovine SP-A, SP-B, and SP-C.

In the present study, we isolated complete cDNAs for ovine SP-A and SP-C and two cDNA forms for SP-B that differ by a deletion of 69 bp within the sequence encoding the active airway protein. Until now, several studies have identified mutations and polymorphisms within the human SP-B gene (reviewed in Ref. 13). Different isoforms of SP-B transcripts, which are thought to be the result of either genomic mutations or abnormalities in RNA splicing, are known (32). Alternative splicing was reported in the 3'-untranslated region of rabbit SP-B (27). In the mouse, rat, and rabbit, the deletion within the region of the mature SP-B protein that matches the deletion we have identified within one of our ovine SP-B cDNA clones was recently detected by PCR.

MATERIALS AND METHODS

Poly(A)+ RNA extraction. Lung biopsies were obtained from anesthetized healthy adult sheep, immediately immersed into liquid nitrogen, and homogenized to isolate poly(A)+ RNA with PolyATtract System 1000 (Promega, Mannheim, Germany). Briefly, ~0.5 g of lung tissue was removed from...
Fig. 1A. Nucleotide and deduced amino acid sequences of ovine surfactant protein (SP) A cDNA. Nucleotides are numbered from 5′ to 3′, with nucleotide 1 representing 1st base of initiation codon. Amino acids are numbered 1–248 beginning with Met. Nucleotide sequence contains an open reading frame of 744 nucleotides, predicting a protein of 248 amino acids. Boldface type, location of putative collagen-like region of 73 amino acids with 24 Gly-X-Y repeats; open letters, consensus sequence for potential N-linked oligosaccharide addition at Asn207; starred residues that can form disulfide pairing; underlined type, 6 ATTTA elements in 3′-untranslated region; boldface and underlined type, poly(A) addition signal AATAAA.
liquid nitrogen, minced on ice, and homogenized in 4 ml of 4 M guanidine thiocyanate-25 mM sodium citrate-2% β-mercaptoethanol. A total of 500 pmol of biotinylated oligo(dT) probe was added to the homogenate and incubated at 70°C for 5 min. The mixture was centrifuged at 12,000 g for 10 min at room temperature, and the supernatant was mixed with Streptavidin MagneSphere Paramagnetic Particles (SA-PMPs; Promega) in 6 ml of 0.5× saline-sodium citrate (SSC; 1× SSC is 150 mM NaCl and 15 mM trisodium citrate, pH 7.0) at room temperature for 5 min. The SA-PMPs were magnetically separated and washed three times with 0.5× SSC. The poly(A)+ RNA was eluted with nuclease-free water and quantified by measuring absorption at 260 nm. The poly(A)+ RNA was concentrated by precipitation.

cDNA library construction. An ovine cDNA library from lung tissue was constructed in λ ZAP II (Stratagene, Heidelberg, Germany). The method comprises reverse transcription of poly(A)+ RNA by Moloney murine leukemia virus reverse transcriptase with a mixture of random primers and an oligo(dT) linker-primer containing an XhoI restriction endonuclease site. After ligation of EcoR I adapters and digestion with XhoI, cDNA was inserted into the λ ZAP II phagemid unidirectionally. The library contained 3 × 10^6 independent clones, the size of which exceeded 400 bp on average.

Library screening. About 1,000,000 plaque-forming units of the amplified library were plated on 20 plates that had a diameter of 150 mm. The phages were transferred to sterile Qiabrané nylon membranes (Qiagen, Hilden, Germany), soaked for 2 min in 1.5 M NaCl-0.5 M NaOH, and neutralized for 5 min in a solution of 1.5 M NaCl and 0.5 M Tris-HCl, pH 8.0. The membranes were rinsed in 0.2 M Tris-HCl, pH 7.4, and 2× SSC and cross-linked by ultraviolet light (1,200 mJ; UV Stratalinker 1800, Stratagene). The membranes were prehybridized in solution containing 2× PIPES buffer, 50% formamide, 0.5% sodium dodecyl sulfate (SDS), and 100 µg/ml of sonicated denatured salmon sperm DNA for at least 2 h.

Table 1. Comparison of deduced amino acid sequences of SP precursors from sheep with sequences from other species

<table>
<thead>
<tr>
<th>Species</th>
<th>Sheep SP-A</th>
<th>Sheep SP-B</th>
<th>Sheep SP-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog</td>
<td>76</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Guinea pig</td>
<td>72</td>
<td>65</td>
<td>80</td>
</tr>
<tr>
<td>Human</td>
<td>74</td>
<td>72</td>
<td>80</td>
</tr>
<tr>
<td>Mink</td>
<td>74</td>
<td>72</td>
<td>82</td>
</tr>
<tr>
<td>Monkey</td>
<td>79</td>
<td>68</td>
<td>84</td>
</tr>
<tr>
<td>Mouse</td>
<td>79</td>
<td>68</td>
<td>79</td>
</tr>
<tr>
<td>Pig</td>
<td>81</td>
<td>66</td>
<td>74</td>
</tr>
<tr>
<td>Rabbit</td>
<td>78</td>
<td>69</td>
<td>74</td>
</tr>
<tr>
<td>Rat</td>
<td>75</td>
<td>67</td>
<td>78</td>
</tr>
</tbody>
</table>

Values are identity in percent. Sequences of sheep (clone SP-B34), dog, and guinea pig surfactant protein (SP) B are incomplete.
Fig. 2A — For legend see facing page.
Fig. 2A. Nucleotide and deduced amino acid sequences of ovine SP-B cDNAs from clone SP-B34 (top sequence) and clone SP-B75 (bottom sequence). Sequence comparison with other species points to fact that this sequence is possibly lacking initiation codon. Therefore, we numbered nucleotides from 5' to 3', with nucleotide 4 representing 1st base of 2nd codon. Amino acids are numbered 1-375. SP-B34 contains an open reading frame of 1,122 nucleotides, predicting a protein of 374 amino acids. Open reading frame of SP-B75 comprises 615 nucleotides, predicting a protein of 205 amino acids. Boldface type, active airway proteins of 79 amino acids for SP-B34 and 56 amino acids for SP-B75; *, Cys candidates that form intrachain disulfide bridges; **, Cys that forms an interchain disulfide link; open letters, 2 potential sites of N-linked glycosylation at Asn75 and Asn305.
4 h at 42°C, hybridized with 1 × 10^6 counts·min^{-1} (cpm)·ml probe in fresh hybridization solution for 16 h at 45°C, and washed in 0.1× SSC-0.1% SDS at 60°C with shaking. The membranes were exposed to Kodak BioMax MS film with a Kodak MS intensifying screen (Kodak, Cambridge, UK) for 4 h at 280°C. Positive phages were purified to homogeneity by replating at 10- and 100-fold lower phage dilutions and rescreening. Pure plaque clones were subjected to in vivo excision of the pBK-CMV phagemid from the λZAP II phage according to the manufacturer’s instructions (Stratagene).

Southern blot. To identify clones containing SP-A, SP-B, or SP-C sequences, DNA of the positive clones was digested with...
restriction endonucleases and characterized by Southern hybridization. The DNA was fractionated on a 1% agarose gel, denatured with 1 M NaCl and 0.5 M NaOH two times for 15 min each, neutralized with a buffer containing 0.5 M Tris·HCl, pH 7.4, and 1.5 M NaCl two times for 15 min each, and transferred to Zeta-Probe nylon membranes (Bio-Rad, Munich, Germany) with 10^3 SSC. After ultraviolet cross-linking was performed, the membranes were prehybridized for 2 h at 45°C with 6× SSPE-0.5% SDS-50% formamide-50 µg/ml denatured salmon sperm DNA. The membranes were hybridized in 6× SSPE-0.5% SDS-50% formamide-50 µg/ml denatured salmon sperm DNA for 16 h at 45°C with the same 32P-labeled probes as used for the cDNA library screening. The membranes were rinsed two times each in 6× SSPE-0.1% SDS at room temperature, 1× SSPE-0.5% SDS at 37°C, and 0.1× SSPE-0.5% SDS at 65°C and subjected to autoradiography.

Sequencing and analysis. Sequencing of double-stranded DNA was done on a Pharmacia (Freiburg, Germany) ALF system automated DNA sequencer with internal labeling with fluorescein dATP produced with the AutoRead sequencing kit (Pharmacia). T3 and T7 promoter primers were used to prime the sense and antisense strands, and then additional oligonucleotides were synthesized from previously sequenced regions and used as primers to extend the nucleotide sequences. All cDNA clones were sequenced on both strands. Sequence analysis and alignment were performed on an Apple Macintosh computer with GeneWorks 2.3.1 software. For amino acid sequence comparisons between species, the following conditions were applied: if there were different sequences for one species, the first published sequence that was complete was selected for alignment.

RNase protection assay. We subcloned an EcoRI-EcoRV fragment of 282 bp from the clone SP-B34 into the vector pGEM-3Z (see Fig. 3A). The resulting construct was linearized with EcoRI to generate an antisense riboprobe with RNA polymerase SP6 essentially according to the instructions for the Riboprobe Systems Kit of Promega. The RNase protection assay (RPA) was performed with the RPA II kit (Ambion, Austin, TX) following the manufacturer’s instructions. Total RNA (1.25–10 µg) from whole lung cells was applied for the assay. About 3×10^7 cpm (1 µl) of the labeled SP-B antisense probe were combined with 20 µl of hybridization buffer (80% deionized formamide, 100 mM sodium citrate, pH 6.4, 300 mM sodium acetate, pH 6.4, and 1 mM EDTA) and hybridized for 18 h at 42°C. RNase digestion was performed with 11 U of RNase T1 and 0.05 U of RNase A. After an incubation period of 30 min at 15°C, the reaction was stopped and precipitated by the addition of 300 µl of inactivation solution (RPA II kit, Ambion). For sequencing, the protected fragments were run on a 5% polyacrylamide gel in the presence of 8 M urea and 1× Tris-borate-EDTA (0.9 M Tris·HCl, 0.9 M boric acid, and 20 mM EDTA, pH 8.0) for 1 h at 250 V. The gel was transferred to Whatman 3MM filter paper (Whatman, Maidstone, UK) and exposed to NEN Reflection autoradiography film (NEN Life Science Products, Cologne, Germany) at room temperature. For syn-
thesis of labeled RNA molecular-size standards, we used the Century Marker Template sets (Ambion), a mixture of linearized plasmids ready for use as templates for in vitro transcription reactions.

Probe preparation. Human cDNA probes specific for SP-A, SP-B, and SP-C (kindly provided by J. Floros, Pennsylvania State University, Hershey, PA) were used in situ hybridization to screen the cDNA library. The probe specific for SP-A had a length of 0.93 kb, that for SP-B of 2.0 kb, and that for SP-C of 0.85 kb. The probes were labeled by random hexamer priming (random-primer DNA labeling kit, Boehringer Mannheim, Mannheim, Germany) with $[^{32P}]dCTP$ (3,000 Ci/mmol) to a specific activity of $10^8$ to $10^9$ cpm/µg. Unincorporated nucleotides were removed with NucTrap push columns (Stratagene). To enrich for SP-B clones containing NH$_2$-terminal sequences, we subcloned a 500-bp EcoR I-Sma I fragment of human SP-B cDNA into the vector pGEM-4Z. The resulting subclone, SP-B5, comprises the 5’-end of the human SP-B cDNA sequence. For further screening of the cDNA library, SP-B5 was used to generate an $[^{32P}]$CTP (800 Ci/mmol)-labeled riboprobe with a specific activity of $5 \times 10^8$ cpm/µg with the Riboprobe systems kit (Promega).

RESULTS

Isolation of SP-A, SP-B, and SP-C cDNAs. We constructed a cDNA library from ovine lung in λ ZAP II and isolated positive clones using human cDNA probes specific for SP-A, SP-B, and SP-C. We selected strongly positive clones and purified them to homogeneity. We found cDNA inserts for SP-A of 1.9 kb, for SP-B of 0.5–1.7 kb, and for SP-C of 0.8 kb from the positive clones after DNA digestion with EcoRI and XhoI and agarose gel electrophoresis. Southern blot analysis of digested DNA from all phagemid clones proved that the cDNA inserts were specific for SP-A, SP-B, or SP-C.

Sequence of SP-A cDNA. Two positive SP-A clones were isolated, one of which encodes the complete cDNA for ovine SP-A. This SP-A cDNA comprises 1,901 bp, consisting of an open reading frame of 744 nucleotides and 5’- and 3’-untranslated regions of 45 and 1,122 nucleotides, respectively (Fig. 1A). The 3’-untranslated region contains six ATTATA motifs. The polyadenylation signal AAATAAA precedes a long poly(A) tail of 53 As. The nucleotide sequence allows prediction of the primary amino acid sequence of the SP-A precursor protein comprising 248 amino acids. The NH$_2$-terminal region of ovine SP-A includes a putative signal peptide of 20 amino acids followed by a seven-amino acid segment with a single cysteine (Cys$^{24}$). This cysteine is conserved in the SP-A of all species because it is essential for the interchain disulfide bond formation that leads to multimerization of SP-A. Amino acids 28–100 are supposed to represent the collagen-like region. This part of the SP-A sequence is arranged into 24 repeating tripeptides of the sequence Gly-X-Y, where X and Y are any amino acids. The continuity of the collagen-like sequence is broken by the insertion of an extra Gly$^{62}$. After this break, the tripeptides continue for additional 12 repeats. After the collagen-like sequence, the presumed collagenase-resistant sequence of 148 amino acids extends through the COOH-terminal Phe. At Asn$^{207}$ in the COOH-terminal domain, there is the consensus sequence Asn-Tyr-Thr, a known target for N-linked glycosylation, which is highly conserved between SP-A proteins of all species. Furthermore, the cysteine residues are known to form intrachain disulfide bonds, linking together Cys$^{35}$ to Cys$^{246}$ and Cys$^{224}$ to Cys$^{238}$, the locations of which are conserved among many lectins. Alignment of the deduced ovine SP-A precursor protein showed high homology to SP-A sequences from human (53), pig (GenBank accession no. M11769; J. E. Adamou and N.A. Elshourbagy, unpublished data), dog (2), mouse (26), rat (11), guinea pig (55), and rabbit (3), with 81% identity with pig and an average identity of 76% across species (Fig. 1B, Table 1).

Sequence of SP-B cDNAs. The cDNA library was screened first with the complete human SP-B cDNA and second with a subcloned probe specific for the NH$_2$ terminus of SP-B. With this strategy, we could isolate >20 individual SP-B cDNA clones, which differed in length and sequence. Many SP-B clones were lacking ~500 bp of the NH$_2$-terminal sequence. After screening the library with a probe specific for the NH$_2$ terminus, we could isolate SP-B cDNA clones containing NH$_2$-terminal sequences. None of these clones, however, comprise the initiation codon. Two clones, SP-B34 and SP-B75, were fully sequenced and showed lengths of 1,660 and 705 bp.

The SP-B34 cDNA clone does not encode the initiating Met but from comparison with SP-B cDNA sequences of other species, the ovine SP-B precursor protein most likely contains one additional residue upstream from the first codon Ala. An open reading frame of 1,122 nucleotides in SP-B34 predicts a precursor protein of 374 amino acids. The 3’-untranslated region encompasses 538 bp without any polyadenylation recognition site and extends to a poly(A) of 18 As (Fig. 2A).

SP-B34 contains two potential sites of N-linked glycosylation, a conserved one at Asn$^{305}$ in the COOH-terminal region and one in the NH$_2$-terminal region at Asn$^{75}$. Within the sequence coding for the active airway protein, six conserved cysteines are expected to form intrachain disulfide bonds, the actual pairing being as follows: from Cys$^{8}$ to Cys$^{27}$, from Cys$^{11}$ to Cys$^{71}$, and from Cys$^{35}$ to Cys$^{56}$. In addition, an intrachain disulfide bond at Cys$^{48}$ is capable of linking two SP-B proteins to form a dimer.

The SP-B75 cDNA clone is partial; it spans the region encoding the active airway protein and extends into the NH$_2$-terminal region. There is an open reading frame of 615 nucleotides, predicting a precursor protein of 205 amino acids. Clone SP-B75 differs from SP-B34 in two ways. The first difference is a deletion of 69 bp in the sequence encoding the active airway protein, with no change in reading frame. The second difference is an exchange of two base pairs, G for A at position 704, of the SP-B75 sequence. The mature SP-B75 protein should be able to form intrachain disulfide bridges from Cys$^{8}$ to Cys$^{27}$ and from Cys$^{11}$ to Cys$^{71}$. It lacks Cys$^{35}$ and Cys$^{48}$, whereas Cys$^{48}$ is present.

The RPA yielded two protected fragments of 282 and 196 bp, which represented the predicted sizes for the
normal SP-B34 sequence and for the sequence of the deletion clone SP-B75, respectively (Fig. 3). With the assumption that the SP-B75 transcript extends further toward the 5′-end than the corresponding cDNA, the protected fragments of this short isoform should comprise 198 and 15 bp. However, the 15-bp fragment of the short isoform could not be resolved because it comigrates with residual degradation products of the RNases at the gel bottom.

Comparison of the SP-B precursor protein from sheep with the respective protein from human (22), dog (17), rat (9), mouse (4), guinea pig (GenBank accession no. AF033190; C. D. Bingle, H. T. Yuan, and S. Gowan, unpublished data), and rabbit (54) shows that an amino acid homology of ~70% is maintained across species (Fig. 2B, Table 1).

Sequence of SP-C cDNA. Screening for ovine SP-C yielded 25 positive hybridization signals from which we isolated a complete SP-C cDNA. The overall length of ovine SP-C cDNA is 809 bp, with an open reading frame of 570 bp encoding a primary translation product of 190 amino acids as shown in Fig. 4A. There is a 5′-untranslated region of 12 bp and a 3′-untranslated region of 227 bp containing the cryptic polyadenylation site ATTAAA at position 763. Similar to SP-A and SP-B, it is also commonly accepted that for SP-C the active airway protein is the result of protein processing at the amino and carboxy termini of a larger precursor. The predicted form of fully processed ovine SP-C is composed of 35 mainly hydrophobic amino acids extending from residues 24 to 58 of the precursor protein. The first residue of the mature ovine SP-C is Leu and the carboxy-terminal residue is also Leu. The NH2-terminal part of the mature protein contains two juxtapositioned cysteines, Cys5 and Cys6, which are known to have one thioester-linked palmitoyl group.

A comparison of the deduced amino acid sequence of ovine SP-C with other vertebrate SP-C precursor proteins is shown in Fig. 4B. Percentages of the identical amino acids (Table 1) indicate a significant homology between the SP-C precursor proteins. Between the SP-C sequences of sheep and monkey (GenBank accession no. U06694; G. An, G. Luo, Y. Zhao, C. Plopper, and R. Wu, unpublished data), mink (GenBank accession no. Z19516; J. Christensen, J. Belusov, T. Storgaard, B. Aasted, and S. Alexandersen, unpublished data), human (16), mouse (15), rat (12), rabbit (GenBank accession no. AF037445; R. K. Margana and V. Boggaram, unpublished data), and dog (47), where only the sequence of the mature protein is known, the average homology is ~80%.

Fig. 3. RNase protection assay for ovine SP-B mRNA. A: plasmid map of SP-B cDNA subclone used to generate an antisense riboprobe with SP6 RNA polymerase. The 323-bp antisense riboprobe protects a region of 282 bp of normal SP-B transcript. Hybridization of riboprobe with deleted transcript (clone SP-B75) yields a protected fragment of 196 bp. B: lanes 1–4, increasing amounts (1.25–10 µg) of total sheep lung RNA digested with RNase A and RNase T1; lanes 5–8, sheep lung RNA digested with varying concentrations of RNase A and/or T1; lane 9, RNA digested solely with RNase T1; lanes 10 and 11, yeast RNA with RNase digestion; lane 12, yeast RNA without RNase digestion, exhibiting full-length probe. In lanes 1–8, predicted protection fragment of 282 bp as well as 196-bp fragment is seen.
Fig. 4A. Nucleotide and deduced amino acid sequences of ovine SP-C cDNA. Nucleotides are numbered from 5' to 3', with nucleotide 1 representing 1st base of initiation codon. Amino acids are numbered 1–190 beginning with Met. Nucleotide sequence contains an open reading frame of 570 nucleotides, predicting a protein of 190 amino acids. The 5'-and 3'-untranslated regions consist of 12 and 227 nucleotides, respectively. Boldface and underlined type, cryptic polyadenylation signal; boldface type, active airway protein of 35 amino acids; *, Cys present as palmitoylcysteine thioesters.
The sequences described have been submitted to the GenBank data library under accession nos. AF076633 for SP-A, AF076636 and AF107544 for SP-B, and AF076634 for SP-C.

**DISCUSSION**

This work provides the cDNAs for ovine SP-A, SP-B, and SP-C and the deduced amino acid sequences of their primary translation products. Our data confirm that the three SP sequences are quite conserved across diverse mammalian species. This suggests that these proteins are subject to highly specific structural constraints, which are essential for their alveolar function. In addition, we found some features that are specific for the ovine sequences. In regard to ovine SP-A, six ATTTA motifs exist within the 3'-untranslated region. In contrast, human SP-A has only one ATTTA motif located near the polyadenylation signal. Such elements within mRNAs have been shown to affect mRNA stability by promoting rapid deadenylation of mRNA and rapid decay characteristic of many unstable growth factor-inducible mRNAs (5, 46). Whether or not the six ATTTA motifs in ovine SP-A have regulatory implications in sheep has to be elucidated.

Concerning ovine SP-B, we isolated and fully sequenced two individual cDNAs, SP-B34 and SP-B75. If we compare the proposed ovine SP-B34 protein to other species, we found that like in humans but unlike in all

<table>
<thead>
<tr>
<th></th>
<th>sheep</th>
<th>monkey</th>
<th>mink</th>
<th>human</th>
<th>mouse</th>
<th>rat</th>
<th>rabbit</th>
<th>dog</th>
</tr>
</thead>
<tbody>
<tr>
<td>sheep</td>
<td>MDVGSKEVLM</td>
<td>ESPPDYSAVP</td>
<td>GGLRIPCCP</td>
<td>VNIRLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monkey</td>
<td>MDVGSKEVLM</td>
<td>ESPPDYSAAP</td>
<td>RGFRGIPCCP</td>
<td>VHLKLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mink</td>
<td>MDVGSKEVLI</td>
<td>ENPPDYSAAP</td>
<td>QGGFLPCFP</td>
<td>SSLKLRIII</td>
<td>IVLIVV87</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human</td>
<td>MDVGSKEVGL</td>
<td>ESPPDYSAAP</td>
<td>RGFRGIPCCP</td>
<td>VHLKLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>MDMSKEVLM</td>
<td>ESPPDYSAAP</td>
<td>RSQFRIPCCP</td>
<td>VHLKLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rat</td>
<td>MDMSGKEVLM</td>
<td>ESPPDYSTGP</td>
<td>RSQFRIPCCP</td>
<td>VHLKLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit</td>
<td>MDMSGKEALM</td>
<td>ESPPDYSAAP</td>
<td>RGFRGIPCCP</td>
<td>VHLKLLIVV</td>
<td>VVVLVWWW</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dog</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>sheep</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSI-AGPEAQ</td>
<td>QRLALSERVG</td>
<td>TTATFSIGST</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monkey</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSI-GAPEAQ</td>
<td>QHLARSGHLV</td>
<td>TTATFSFST</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mink</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSI-AGPEAQ</td>
<td>QRLALSERVG</td>
<td>TTATFSIGST</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSI-AGPEAQ</td>
<td>QRLALSERVG</td>
<td>TTATFSIGST</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSIGAPETQ</td>
<td>KRLAIDEO</td>
<td>TATFSSTG</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rat</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSIGAPETQ</td>
<td>KRLAIDEO</td>
<td>TATFSSTG</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit</td>
<td>VGALLMLHLM</td>
<td>SQQHTEMVLE</td>
<td>MSI-GAPEVQ</td>
<td>QRLALSEWAG</td>
<td>TTATFSIGST</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dog</td>
<td>VGALLMLG</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>sheep</td>
<td>GTVVVDYQRL</td>
<td>LIAYKPAFGT</td>
<td>CCYICMKVAPQ</td>
<td>SIPSLEALTR</td>
<td>KLNPFOAKPP</td>
<td>149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monkey</td>
<td>GLVYDQRL</td>
<td>LIAYKPAFQT</td>
<td>WCYIMKTAPE</td>
<td>SIPSLEALTR</td>
<td>KVPNFAQAKPA</td>
<td>149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mink</td>
<td>GIVYDQRL</td>
<td>LIAYKPAFQT</td>
<td>CCYICMKVAPQ</td>
<td>SIPSLEALTR</td>
<td>KFONFAQKPA</td>
<td>149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human</td>
<td>GLVYDQRL</td>
<td>LIAYKPAFQT</td>
<td>CCYICMKVAPQ</td>
<td>SIPSLEALTR</td>
<td>KFONFAQKPA</td>
<td>149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>GIVYDQRL</td>
<td>LTAYKPAGT</td>
<td>YCYIMKTAPE</td>
<td>SIPSLEAFAK</td>
<td>KLQNFRAKPS</td>
<td>149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rat</td>
<td>GIVYDQRL</td>
<td>LTAYKPAGT</td>
<td>YCYIMKTAPE</td>
<td>SIPSLEALAR</td>
<td>KFONFAQKSS</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit</td>
<td>GIVYDQRL</td>
<td>LTAYKPAGT</td>
<td>YCYIMKTAPE</td>
<td>SIPSLEALAR</td>
<td>KFONFAQKSS</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dog</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>sheep</td>
<td>VPSKLGQEQ</td>
<td>GRDAGSXF-S</td>
<td>---G-DLAFL</td>
<td>GRTVSTLCGE</td>
<td>VPLYYT</td>
<td>190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monkey</td>
<td>VPTSKLDQVE</td>
<td>GRDAGSXF-S</td>
<td>---G-DLAFL</td>
<td>GMAVSTLCGE</td>
<td>VPLYI</td>
<td>191</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mink</td>
<td>VSTSKLGQVE</td>
<td>GHNASAS-P</td>
<td>---G-DLDFL</td>
<td>GTVSTLCGE</td>
<td>VPLYI</td>
<td>190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human</td>
<td>VPTSKLGQAE</td>
<td>GRDAGSXF-S</td>
<td>---G-DPFL</td>
<td>GMAMVTLCGE</td>
<td>VPLYI</td>
<td>191</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>TPSKLGQEE</td>
<td>GHDTGQSESD</td>
<td>---SGRDLAFL</td>
<td>GLAVSTLCGE</td>
<td>LPLYI</td>
<td>193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rat</td>
<td>TPSKLGQEE</td>
<td>GHDSQSDESD</td>
<td>---SGRDLAFL</td>
<td>GLAVSTLCGE</td>
<td>LPLYI</td>
<td>194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit</td>
<td>EPPT-----QR</td>
<td>QCDKGAAPGP</td>
<td>ASSGELAFL</td>
<td>GAVVSTLCGE</td>
<td>VPLYI</td>
<td>188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dog</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>

Fig. 4B. Alignment of SP-C protein predicted by cDNA from sheep compared with that from 7 other species. Shaded boxes, amino acids conserved between all species; hyphens, gaps that were inserted to facilitate alignment; solid lines, amino acids of putative active protein.
other species, both flanking regions of the ovine protein have potential glycosylation sites. ProSP-B34 contains a conserved site of N-linked glycosylation at Asn\(^{205}\) in the COOH-terminal region, which is present in all SP-B sequences, and another site, Asn-Ile-Ser, in the NH\(_2\)-terminal region at Asn\(^{75}\). In human SP-B, the second glycosylation site, Asn-Gln-Thr, is present at Asn\(^{129}\).

The occurrence of more than one type of SP-B cDNA in sheep suggests that more than one form of mRNA encodes SP-B. There are several reports of SP-B transcript deviations. In humans, a frameshift mutation in codon 121 of the SP-B cDNA that has been associated with hereditary SP-B deficiency results in premature termination of translation. In addition to this mutation, one affected patient showed the presence of two transcripts, one having exon 7 deleted and the second with exon 7 being present but using a cryptic splice site in exon 8. The authors (32) suggest that abnormalities of RNA splicing give rise to the two aberrant transcripts. In the rabbit, two isoforms of SP-B were found that differ by an insertion of 69 bp in the 3′-untranslated region (27). In the mouse, rat, and rabbit but not in human, two SP-B mRNAs were detected by PCR, one of which revealed a 69-bp deletion at the same position as the deleted isoform of ovine SP-B in our study (6). In mouse, rat, and rabbit, the different isoforms of SP-B were proposed to result from alternative splicing. The deletion in our ovine SP-B75 is located within the putative exon 6 at the 5′ boundary and therefore could arise from alternative splicing as well.

Comparison of the two cDNAs (SP-B34 and SP-B75) shows that they differ in length, by 2-bp mismatches at the 3′-end, and by a 69-bp deletion in the region encoding the active airway protein. The different lengths of these two cDNAs are due to the fact that one of the two clones, SP-B75, is only partial. The two mismatches between the ovine SP-B sequences SP-B75 and SP-B34 indicate that the two sequences are not identical at the 3′-end. The 69-bp deletion in the region coding for the active airway protein in one of the two cDNAs has functional implications because it abolishes the disulfide link between Cys\(^{35}\) and Cys\(^{46}\). The complexity of differences between the two SP-B clones might well be due to two different genes encoding SP-B. Such a mechanism was demonstrated for human SP-A. More than a decade ago, the first human SP-A gene (SP-A1) was reported (53). When Floros et al. (14) cloned two distinct SP-A cDNA sequences, the existence of two very similar but distinct SP-A genes in humans was likely and finally proved by isolation of a second SP-A gene (24) that corresponded to one of the cDNAs described by Floros et al. (14). It might be species dependent whether there are one or more genes for one or the other SP. Although human, and only human, SP-A is encoded by two genes, for SP-B, only one gene has been identified so far (36). Based on our finding of two distinct ovine cDNAs for SP-B, two genes might well exist in sheep. However, this speculation has to be proved by Southern blot analysis.

With the assumption that the deleted SP-B transcript is translated, as suggested by Chi et al. (6) for the short murine SP-B isoform, the missing disulfide link from Cys\(^{35}\) to Cys\(^{46}\) in SP-B75 would result in major structural deviation. The corresponding functional implications are likely to be substantial. Support for this view comes from a study (1) in transgenic mice, from a study (30) in humans where mutations in the SP-B gene are associated with severe lung disease, and from findings (33) of mutated forms of saposins, a family of proteins that is sequence related to SP-B.

In transgenic mice, the ablation of the cysteine bridge by mutation of Cys\(^{35/46}\) → Ser did not affect intracellular trafficking in vitro but was associated with respiratory failure in vivo. Increased neonatal mortality was observed when both the normal SP-B and the mutant form of SP-B were expressed in transgenic mice and expression of the mutant SP-B alone was lethal (1). This study implicates the Cys\(^{35}\) to Cys\(^{46}\) bond to be critical for SP-B function, and the authors suggest that the formation of SP-B heterodimers may be lethal even in the wild-type background. In our study, the sheep that served for the cDNA library had no respiratory failure. Although both studies support the view for functional importance of the Cys\(^{35}\) to Cys\(^{46}\) bridge, only the study in transgenic mice demonstrated respiratory failure caused by mutant SP-B. Our data from sheep suggest that both forms of SP-B, if expressed, do not cause respiratory disease. However, we do not know yet whether species differences or regulatory mechanisms determine the expression of one or the other or both types of SP-B in sheep.

In addition to data on structural modifications of SP-B with functional consequences in transgenic mice, the alteration of SP-B structure in humans has an impact on SP-B function. Mutations in the SP-B gene have been shown to cause severe lung disease (30), and polymorphisms in the SP-B gene may be associated with the development of respiratory distress in premature infants and adults (13, 34).

More support for the importance of the disulfide structure for the functional properties of SP-B protein comes from members of a family of sequence-related proteins, the saposins. These proteins share strictly conserved structural features with SP-B, with implications for helix topology and lipid interaction (33). The similarities between saposins and SP-B include six cysteine residues with the same disulfide bonding pattern. The saposins are cofactors for various glycolipid hydrolyses that act on sphingolipids (25). From in vitro observations on mutated saposins, it is apparent that the action of several lysosomal hydrolases is altered if the saposin molecule fails to form normal disulfide bonds (39, 50). The importance of the disulfide bridges in maintaining the structure of the protein is stressed by the ability of mutations that disrupt one of these bonds, by changing a cysteine to another amino acid, to interfere with function (19, 20, 40, 45). For example, in two patients with Gaucher disease, mutations have been found in the saposin C gene that result in Cys\(^{72}\) substituted with either Phe or Gly (40, 45). Glucosyle-
ramide accumulates in the patients' tissue despite the presence of normal glucosylceramidase, indicating that the Cys50 to Cys52 disulfide bridge is essential for the saposin C cofactor function.

With regard to ovine SP-C, its sequence differs from other SP-C sequences because the last residue of the SP-C precursor protein is Thr, whereas the last residue in SP-C of all other species is Ile. Furthermore, the first residue of mature ovine and canine SP-C is Leu, whereas other SP-C sequences start with Phe. From the three SP genes, the SP-C gene is most conserved, and mutations in the SP-C gene have not yet been identified or shown to cause lung disease. However, given the apparent importance of SP-C in surfactant function, alterations in the SP-C gene remain a possibility.

In this study, we reported the ovine sequences for SP-A, SP-B, and SP-C. The discovery of a second SP-B cDNA with a 69-bp deletion might be due to differential splicing or to a second SP-B gene in sheep. Elucidating the biological significance of the different SP-B forms will be a challenging and fascinating problem for future research.

We thank D. Pistol for skillful technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft.

Part of this work was presented as a poster at the 5th Marburg Surfactant Symposium, Surfactant and Alveolar Biology, Phillips University, Marburg, Germany, September 1998. Address for reprint requests and other correspondence: U. Pison, Medizinische Fakultät der Humboldt-Universität Berlin, Charité Campus Virchow-Klinikum, Klinik für Anaesthesiologie und operative Intensivmedizin, Augustenburger Platz 1, D-13353 Berlin, Germany (E-mail: ulrich.pison@charite.de).

Received 3 J une 1999; accepted in final form 5 August 1999.

REFERENCES


27. Margana RK and Boggaram V. Transcription and mRNA stability regulate developmental and hormonal expression of...